

# Variational Retrieval of Rain Microphysics and Related Parameters from Polarimetric Radar Data with a Parameterized Operator

VIVEK N. MAHALE AND GUIFU ZHANG

*School of Meteorology, and Advanced Radar Research Center, and Center for Analysis and Prediction of Storms,  
University of Oklahoma, Norman, Oklahoma*

MING XUE

*School of Meteorology, and Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, Oklahoma*

JIDONG GAO

*NOAA/National Severe Storms Laboratory, Norman, Oklahoma*

HEATHER D. REEVES

*NOAA/National Severe Storms Laboratory, and Cooperative Institute for Mesoscale Meteorological Studies,  
University of Oklahoma, Norman, Oklahoma*

(Manuscript received 18 November 2018, in final form 10 September 2019)

## ABSTRACT

A variational retrieval of rain microphysics from polarimetric radar data (PRD) has been developed through the use of S-band parameterized polarimetric observation operators. Polarimetric observations allow for the optimal retrieval of cloud and precipitation microphysics for weather quantification and data assimilation for convective-scale numerical weather prediction (NWP) by linking PRD to physical parameters. Rain polarimetric observation operators for reflectivity  $Z_H$ , differential reflectivity  $Z_{DR}$ , and specific differential phase  $K_{DP}$  were derived for S-band PRD using T-matrix scattering amplitudes. These observation operators link the PRD to the physical parameters of water content  $W$  and mass-/volume-weighted diameter  $D_m$  for rain, which can be used to calculate other microphysical information. The S-band observation operators were tested using a 1D variational retrieval that uses the (nonlinear) Gauss–Newton method to iteratively minimize the cost function to find an optimal estimate of  $D_m$  and  $W$  separately for each azimuth of radar data, which can be applied to a plan position indicator (PPI) radar scan (i.e., a single elevation). Experiments on two-dimensional video disdrometer (2DVD) data demonstrated the advantages of including  $\Phi_{DP}$  observations and using the nonlinear solution rather than the (linear) optimal interpolation (OI) solution. PRD collected by the Norman, Oklahoma (KOUN) WSR-88D on 15 June 2011 were used to successfully test the retrieval method on radar data. The successful variational retrieval from the 2DVD and the radar data demonstrate the utility of the proposed method.

## 1. Introduction

Observation-based radar retrievals utilize radar data to retrieve rain microphysics information. Basic retrievals use empirical formulas to calculate rain microphysics information on a gate-by-gate basis (e.g.,  $Z$ – $R$  relationship for rain rate) without accounting for statistics of the observation errors. Advanced techniques such as optimal interpolation (OI) (Eliassen 1954; Gandin 1963) and

variational methods (Lorenc 1986), which have been used for data assimilation (DA) in numerical weather prediction (NWP), can be utilized in observation-based retrievals to account for these statistics. Early studies on radar observation-based retrievals using variational methods focused on retrievals of three-dimensional wind fields from Doppler velocity observations (Sun et al. 1991; Qiu and Xu 1992; Laroche and Zawadzki 1994; Gao et al. 1999; Gao et al. 2001). Hogan (2007), Cao et al. (2013), and Yoshikawa et al. (2014) used radar hydrometer related observations to optimally retrieve

---

*Corresponding author:* Vivek N. Mahale, vmahale@ou.edu

DOI: 10.1175/JTECH-D-18-0212.1

© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](https://www.ametsoc.org/PUBSReuseLicenses) ([www.ametsoc.org/PUBSReuseLicenses](https://www.ametsoc.org/PUBSReuseLicenses)).

microphysics information by accounting for background and/or observation errors (i.e., spatial covariance is taken into account).

A variational method requires forward observation operators. A forward observation operator (or forward model) is a transformation, based on physical laws, which converts NWP model state variables to observations (Kalnay 2003). Rodgers (2000) states that “the heart of a successful and accurate retrieval method is the forward model.” Throughout the literature, the forward observation operators are simply called observation operators. The best observation operators must include relevant physics, are numerically efficient, and have easily calculated first derivatives (i.e., the Jacobian) (Rodgers 2000).

An observation-based radar retrieval is beneficial if it can be easily used in convective scale NWP to improve model microphysics parameterization and weather forecasts. Single moment and double moment microphysics parameterization schemes are commonly used in high-resolution convective scale NWP models (e.g., Lin et al. 1983; Milbrandt and Yau 2005; Morrison et al. 2005). In convective-scale NWP model simulations with single-moment microphysics, rainwater mixing ratio  $q_r$  is the only prognostic variable for hydrometeor physics. In double-moment microphysical parameterization schemes, both  $q_r$  and the number concentration  $N_t$  are predicted. The values of  $N_t$  and  $q_r$  can be converted to water content ( $W = \rho_a q_r$ , where  $\rho_a$  is air density) and mass-/volume-weighted diameter  $D_m$ , which are related to the drop size distribution (DSD) for rain (i.e.,  $W$  and  $D_m$  can be considered “related parameters” to describe rain microphysics). They are also considered “physical parameters” because they describe physical quantities and are more likely to have a Gaussian distribution than nonphysical quantities. They can be utilized to calculate other microphysical quantities such as rainfall rate. This motivates us to use parameterized polarimetric forward observation operators that link NWP model state variables and radar observed quantities for an observation-based retrieval. Essentially, the goal is to link the prognostic physics variables of  $N_t$  and  $q_r$  to the polarimetric radar variables by retrieving  $W$  and  $D_m$  for rain.

In this study, an observation-based variational retrieval of rain microphysics from polarimetric radar data (PRD) has been developed through the use of derived parameterized polarimetric observation operators and a nonlinear, iterative solution.

The variational retrieval and derived observation operators are tested by conducting experiments on two-dimensional video disdrometer (2DVD; Kruger and Krajewski 2002) data and radar data. Using 2DVD data, the intrinsic or true values of the retrieved variables are known and simulated observations can be calculated.

Since the truth is known, the 2DVD experiments can quantitatively evaluate the variational retrieval and observation operators for different conditions. These experiments will be discussed in context of the both the retrieval of  $W$  and  $D_m$  (i.e., the inverse problem) and the estimation of polarimetric radar variables (i.e., the forward problem) to test the observation operators. In other words, the observation operators are tested both going backward from the observations to the state variables and forward from the state variables to observations.

Once the variational retrieval and observation operators are successfully tested on 2DVD data and simulated observations, the next step is to apply them to real radar data. The goal of the radar data experiments is to demonstrate the variational retrieval on a both a single azimuth and an entire elevation scan of PRD.

This manuscript is organized as follows. The derivation of the parameterized polarimetric observation operators is shown in section 2. This is followed by a description of the variational retrieval methodology in section 3. Next, the polarimetric observation operators and variational retrieval methodology are applied to DSD data collected by a disdrometer in section 4 and to radar data in section 5. A summary of results, conclusions, and future work for this study are provided in the final section.

## 2. Parameterized polarimetric radar forward observation operators

The DSD is the fundamental description of rain microphysics (Zhang 2016). Therefore, a DSD model will be utilized to derive the observation operators. The DSD is defined as the number of drops in a unit volume for each unit diameter bin. It is a function of equivalent drop diameter  $D$  and is expressed as  $N(D)$  ( $\text{m}^{-3} \text{mm}^{-1}$ );  $D$  is the diameter of a sphere that has an equivalent volume to an oblate spheroid. The raindrops are assumed to be spheroids that become more oblate as their size increases. In this study, the two-parameter exponential distribution will be used for the derivation of the observation operators:

$$N(D) = N_0 \exp(-\Lambda D), \quad (1)$$

where  $N_0$  ( $\text{m}^{-3} \text{mm}^{-1}$ ) is the intercept parameter and  $\Lambda$  ( $\text{mm}^{-1}$ ) is the slope parameter. These are often called “DSD parameters.”

For a two-parameter DSD model, the model physics parameters  $N_t$  and  $q_r$  can be converted to  $W$  and  $D_m$ . Therefore, the state variables chosen for the observation operators are  $W$  and  $D_m$ . Note that  $W$  and  $D_m$  can be derived using existing DSD retrievals that retrieve  $N_0$  and  $\Lambda$ ; however, the advantage is these are physical parameters, which tend to be Gaussian distributed. Nonphysical

parameters are less likely to have a Gaussian distribution. The terms  $N_0$  and  $\Lambda$  do not have a Gaussian distribution (e.g., Cao et al. 2010; Zhang 2016). Cao et al. (2010) showed that  $N_0$  and  $\Lambda$  have greatly skewed distributions.

The first state variable,  $W$  ( $\text{g m}^{-3}$ ), is derived using the definition of the volume of a sphere ( $V = D^3\pi/6$ ) and the third DSD moment:

$$W = \frac{\pi}{6} \rho_w \int_0^{D_{\max}} D^3 N_0 \exp(-\Lambda D) dD \quad (2a)$$

$$= \frac{\pi}{6} \times 10^{-3} N_0 \Lambda^{-4} \gamma(\Lambda D_{\max}, 4), \quad (2b)$$

where  $\rho_w$  is the density of water and  $\gamma$  is the incomplete gamma function. As previously noted,  $W$  is directly related to  $q_r$  as  $W = \rho_a q_r$ .

The second state variable,  $D_m$  (mm), is derived using the ratio of the fourth and third DSD moments:

$$D_m = \frac{\int_0^{D_{\max}} D^4 N_0 \exp(-\Lambda D) dD}{\int_0^{D_{\max}} D^3 N_0 \exp(-\Lambda D) dD} \quad (3a)$$

$$= \frac{1}{\Lambda} \frac{\gamma(\Lambda D_{\max}, 5)}{\gamma(\Lambda D_{\max}, 4)} \quad (3b)$$

Parameter  $N_t$  is related to  $N_0$  by the zeroth DSD moment:

$$N_t = \int_0^{D_{\max}} N_0 \exp(-\Lambda D) dD \quad (4a)$$

$$= \frac{N_0}{\Lambda} \gamma(\Lambda D_{\max}, 1). \quad (4b)$$

The relationship of  $D_m$  to  $q_r$  and  $N_t$  can be derived by first solving for  $N_0$  from (2) and  $\Lambda$  from (4). By substituting  $N_0$  into the equation for  $\Lambda$ , the following equation for  $\Lambda$  is found:

$$\Lambda = \left[ \frac{\frac{\pi}{6} \rho_w N_t \gamma(\Lambda D_{\max}, 4)}{\rho_a q_r \gamma(\Lambda D_{\max}, 1)} \right]^{1/3}. \quad (5)$$

Therefore, using the definition of  $D_m$  from (3),  $D_m$  is related to  $q_r$  and  $N_t$  by the following relationship:

$$D_m = \frac{\gamma(\Lambda D_{\max}, 5)}{\gamma(\Lambda D_{\max}, 4)} \left[ \frac{\rho_a q_r \gamma(\Lambda D_{\max}, 1)}{\frac{\pi}{6} \rho_w N_t \gamma(\Lambda D_{\max}, 4)} \right]^{1/3}. \quad (6)$$

The dual-polarized radar variables used for the derivation of the observation operators are the horizontal radar reflectivity factor  $Z_h$ , differential reflectivity

$[Z_{\text{DR}} = 10 \log_{10}(Z_{\text{dr}})$ , where  $Z_{\text{dr}}$  is in linear units], specific differential phase  $K_{\text{DP}}$ , and copolar correlation coefficient  $\rho_{\text{hv}}$ .

The  $K_{\text{DP}}$  observation operator will be applied through the use of  $\Phi_{\text{DP}}$ , where  $\Phi_{\text{DP}}$  is the total differential phase, which is the sum of the differential backscattering phase and differential propagation phase (or differential phase). It is assumed that contributions of differential backscattering phase are negligible at S band for rain. With the assumption that differential backscattering phase is negligible,  $\Phi_{\text{DP}}$  is assumed to only be an estimate of the differential propagation phase, which is the integral of the specific differential phase over the propagation path. Therefore,  $\Phi_{\text{DP}}$  is directly related to  $K_{\text{DP}}$  by the following:

$$\Phi_{\text{DP}}(r) = 2 \int_0^r K_{\text{DP}}(l) dl. \quad (7)$$

Parameter  $\Phi_{\text{DP}}$  is used in the variational retrieval instead of  $K_{\text{DP}}$  because it is a direct quantity that is measured by radars.

The derivation of the observation operators is a three-step process. The first step is to calculate  $N(D)$  for a range of  $D_m$  using a normalized  $W$  of  $1 \text{ g m}^{-3}$  and assuming an exponential distribution (i.e.,  $Z_h$  and  $K_{\text{DP}}$  are normalized so that  $W$  is  $1 \text{ g m}^{-3}$  and  $D_m$  only varies). To calculate  $N(D)$ ,  $\Lambda$  is first calculated for a range of  $D_m$  using (3). Using  $\Lambda$ ,  $N_0$  can be calculated from (2). Once  $\Lambda$  and  $N_0$  are known,  $N(D)$  can be calculated using (1). For this derivation,  $D_m$  ranges from  $\sim 0.08$  to  $\sim 4.35$  mm when using a finite  $D_{\max}$  of 8 mm.

The second step is to calculate  $Z_h$ ,  $Z_{\text{dr}}$ ,  $K_{\text{DP}}$ , and  $\rho_{\text{hv}}$  by using the S-band scattering amplitudes calculated with the T-matrix method and using the  $N(D)$  values from the first step. The T-matrix method (Waterman 1971) is a numerical-scattering solution. For the scattering amplitude calculations, it is assumed that raindrops will fall with their major axis aligned horizontally (i.e., canting angle mean and standard deviations are zero). It has been found that the true standard deviation for the canting angles for raindrops is somewhere between  $0^\circ$  and  $10^\circ$  (e.g., Bringi and Chandrasekar 2001; Ryzhkov et al. 2002). Ryzhkov et al. (2002) found the assumption of no canting angle only results in a slight overestimation ( $< 6\%$ ) in some PRD, so this is a reasonable assumption. The axis ratios of the raindrops assume the following relationship (Zhang 2016):

$$r = 0.9551 + 0.0251D - 0.03644D^2 + 0.005303D^3 - 0.0002492D^4, \quad (8)$$

where  $r$  is the ratio of the semiminor and semimajor axes.

The third step is to use a polynomial function to fit the calculated  $Z_h$ ,  $Z_{\text{dr}}$ ,  $K_{\text{DP}}$ , and  $\rho_{\text{hv}}$  to  $D_m$ . Because a normalized  $W$  of  $1 \text{ g kg}^{-1}$  was used in the calculation,  $W$

will be multiplied by the polynomial for  $Z_h$  and  $K_{DP}$ . These will cancel for  $Z_{DR}$  and  $\rho_{hv}$  because they are ratios.

As a result, here are the derived parameterized polarimetric radar forward observation operators for rain:

$$Z_h \approx W(-0.3078 + 20.87D_m + 46.04D_m^2 - 6.403D_m^3 + 0.2248D_m^4)^2, \quad (9)$$

$$Z_{dr} \approx 1.019 - 0.1430D_m + 0.3165D_m^2 - 0.06498D_m^3 + 0.004163D_m^4, \quad (10)$$

$$K_{DP} \approx W(0.009260 - 0.08699D_m + 0.1994D_m^2 - 0.02824D_m^3 + 0.001772D_m^4), \quad (11)$$

$$\rho_{hv} \approx 0.9987 + 0.008289D_m - 0.01160D_m^2 + 0.003513D_m^3 - 0.0003187D_m^4, \quad (12)$$

or in logarithmic units for  $Z_H$  and  $Z_{DR}$ :

$$Z_H \approx 10 \log_{10}[W(-0.3078 + 20.87D_m + 46.04D_m^2 - 6.403D_m^3 + 0.2248D_m^4)^2], \quad (13)$$

$$Z_{DR} \approx 10 \log_{10}(1.019 - 0.1430D_m + 0.3165D_m^2 - 0.06498D_m^3 + 0.004163D_m^4). \quad (14)$$

Figure 1 plots all the observation operators from (9)–(14). These observation operators include relevant physics, are numerically efficient (i.e., polynomials), and have easily calculated first derivatives. Note that these calculated polarimetric variables are intrinsic values that do not consider propagation and error effects (e.g., attenuation and sampling errors) and are only valid for S-band radar data.

### 3. Variational retrieval

The solution to an observation-based variational retrieval is an optimal analysis field  $\mathbf{x}_a$  that minimizes a cost function (Lorenc 1986):

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + [\mathbf{y} - H(\mathbf{x})]^T \mathbf{R}^{-1} [\mathbf{y} - H(\mathbf{x})]. \quad (15)$$

The convention here follows vector–matrix form, which is the notation used by the data assimilation community (e.g., Ide et al. 1997; Huang 2000; Kalnay 2003). The uppercase boldface letters represent matrices and the lowercase boldface letters represent vectors in this convention.

The cost function  $J$  is the distance between the analysis  $\mathbf{x}$  and the background  $\mathbf{x}_b$ , weighted by the inverse of the background error covariance  $\mathbf{B}$  plus the distance of the analysis  $\mathbf{x}$  to the observations  $\mathbf{y}$  weighted by the inverse of the observation error covariance  $\mathbf{R}$  (Kalnay 2003).

Parameter  $H$  is the observation operator, and  $\mathbf{x}$  is an analysis vector that is a concatenate of the state variables  $W$  and  $D_m$  (i.e., the variational retrieval finds an optimal analysis of  $W$  and  $D_m$ ). The other terms in the cost function ( $\mathbf{x}_b$ ,  $\mathbf{B}$ ,  $\mathbf{y}$ , and  $\mathbf{R}$ ) are defined in greater detail later in this section.

The solution for an observation-based OI retrieval is (Kalnay 2003)

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{W}[\mathbf{y} - H(\mathbf{x}_b)], \quad (16a)$$

$$\mathbf{W} = \mathbf{B}\mathbf{H}^T(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1}, \quad (16b)$$

where  $\mathbf{W}$  is the optimal weight matrix and  $\mathbf{H}$  is the Jacobian, which is the linear approximation of  $H$ .  $\mathbf{H}$  is derived from (11), (13), and (14), which will be shown later in this section.

The assumption with the OI solution is that background is a reasonable approximation of the true analysis, so that the solution of the analysis is equal to the background values plus small increments (Kalnay 2003). The OI solution is usually not solved iteratively (i.e., it is solved explicitly), and is equivalent to the three-dimensional variational assimilation (3DVAR) solution that is found by minimizing the cost function (15).

For moderately nonlinear problems, the Gauss–Newton iterative method can find the solution (Rodgers 2000):

$$\mathbf{x}_{i+1} = \mathbf{x}_b + \mathbf{W}[\mathbf{y} - H(\mathbf{x}_i) + \mathbf{H}_i(\mathbf{x}_i - \mathbf{x}_b)], \quad (17a)$$

$$\mathbf{W} = \mathbf{B}\mathbf{H}_i^T(\mathbf{R} + \mathbf{H}_i\mathbf{B}\mathbf{H}_i^T)^{-1}, \quad (17b)$$

where  $\mathbf{H}_i$  and  $H(\mathbf{x}_i)$  are updated during each iteration  $i$ .

The Gauss–Newton iterative method is utilized to solve nonlinear least squares problems (e.g., variational data assimilation problems) and has been shown to correspond to the incremental four-dimensional variational (4DVAR) assimilation solution (Lawless et al. 2005a,b). For this study, the Gauss–Newton iterative method, (17), was used to better account for the nonlinearity in the forward operator.

Note that for any variational retrieval, there is an assumption that the background and observation errors have a Gaussian probability distribution, the background and observation errors are uncorrelated, and the background and observations are unbiased (Kalnay 2003). Errors will be introduced in the variational retrieval if these conditions are not met.

A comparison to the OI solution, (16), is also shown in section 4. The comparison between both methods is shown because even though most retrieval problems are not truly linear, many still utilize linearization about some prior state even when the observation operator is not truly linear (Rodgers 2000). Therefore, the OI solution is tested because it provides a baseline to compare to Gauss–Newton iterative solution and to see if linearity can be assumed.

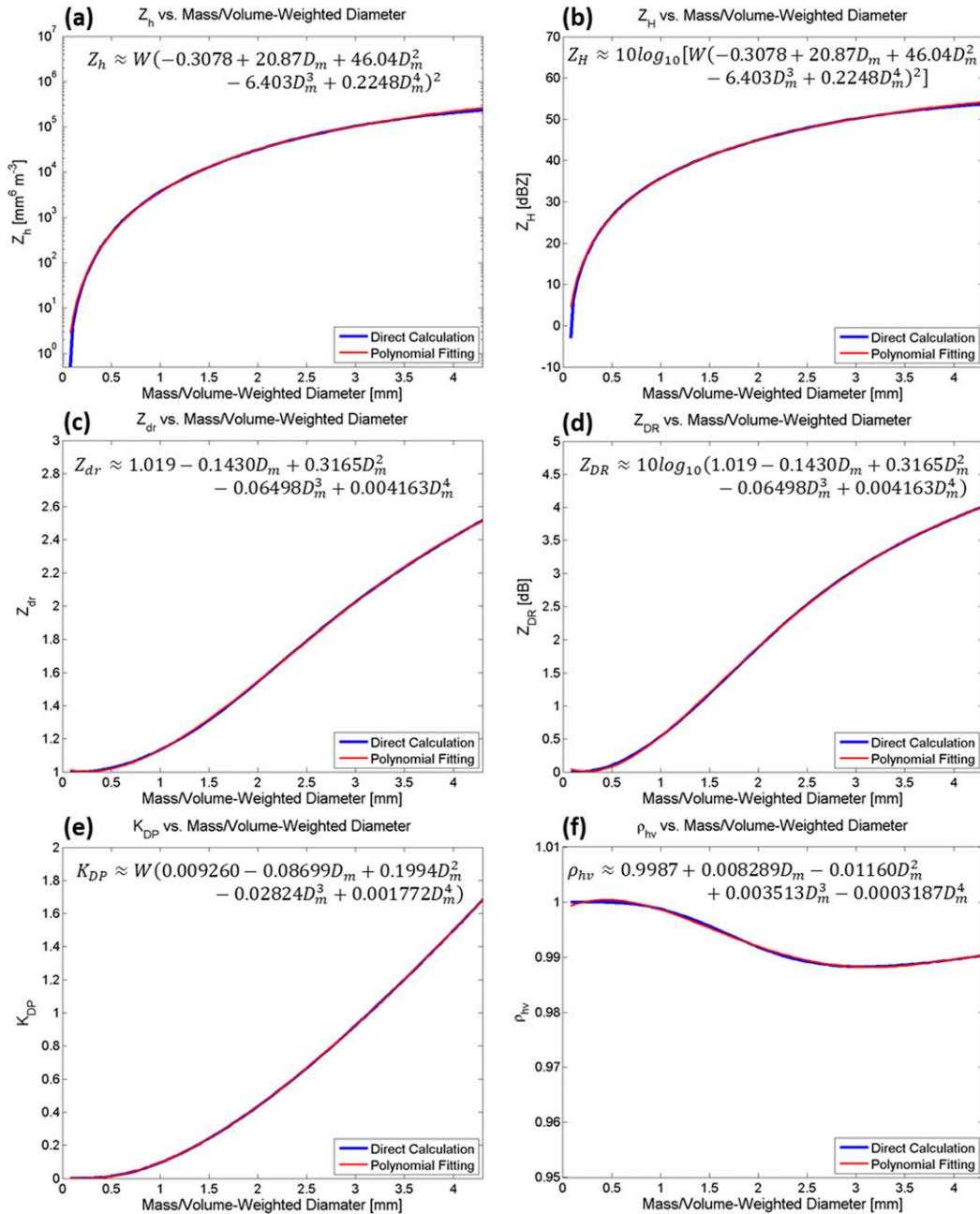


FIG. 1. The fitted parameterized polarimetric forward observation operators compared to direct calculations from T-matrix method for (a)  $Z_h$ , (b)  $Z_H$ , (c)  $Z_{dr}$ , (d)  $Z_{DR}$ , (e)  $K_{DP}$ , and (f)  $\rho_{hv}$ . These are normalized for  $W$  of  $1 \text{ g m}^{-3}$ . The equation in each panel is the derived observation operator.

To solve an observation-based variational retrieval, the state variables and observations must be defined. As previously noted, the state variables for this observation-based retrieval are  $W$  and  $D_m$ , forming the state or analysis vector  $\mathbf{x} = [W; D_m]$ . The observations used are  $Z_H$ ,  $Z_{DR}$ , and  $\Phi_{DP}$ , yielding the observation vector  $\mathbf{y} = [Z_H; Z_{DR}; \Phi_{DP}]$ . Note that  $\rho_{hv}$  is not used because of its limited dynamic range for rain at S-band (Fig. 1f);  $\rho_{hv}$  would not add value

to retrieval for pure rain as its error is approximately equal to its dynamic range.

For the solution to (17), let us look at each term individually. First, assume that it is solved azimuthally and that the number of range gates in one azimuth is  $n_r$ . The variational retrieval is solved azimuthally (1D in the radial direction) to coincide with  $\Phi_{DP}$  measurements (i.e., to coincide with  $\Phi_{DP}$  defined as the integral of  $K_{DP}$  over the

propagation path). Therefore, the dimension of state variables is  $n = 2n_r$  and the dimension of observations is  $p = 3n_r$ .

The analysis vector at iteration  $i$  with dimension  $n \times 1$  ( $2n_r \times 1$ ) is denoted  $\mathbf{x}_i$ . It is the concatenate of the analysis vectors of the state variables  $W$  and  $D_m$  at iteration  $i$ .

The analysis vector at iteration  $i+1$  with dimension  $n \times 1$  ( $2n_r \times 1$ ) is denoted  $\mathbf{x}_{i+1}$ . It is the concatenate of the analysis vectors of the state variables  $W$  and  $D_m$  at iteration  $i+1$ .

The background state vector (also known as the first guess) with dimension  $n \times 1$  ( $2n_r \times 1$ ) is denoted  $\mathbf{x}_b$ . It is the concatenate of the background vectors of the state variables  $W$  and  $D_m$ . For this study, the background state vector is obtained from empirical formulas for  $W$  and  $D_m$  (Zhang 2016):

$$W = (1.023 \times 10^{-3})Z_h \times 10^{-0.0742Z_{DR}^3 + 0.511Z_{DR}^2 - 1.511Z_{DR}}, \quad (18)$$

$$D_m = 0.0657Z_{DR}^3 - 0.332Z_{DR}^2 + 1.090Z_{DR} + 0.689. \quad (19)$$

The observation vector with dimension  $p \times 1$  ( $3n_r \times 1$ ) is denoted  $\mathbf{y}$ . It is the concatenate of the observation vectors for  $Z_H$ ,  $Z_{DR}$ , and  $\Phi_{DP}$ .

The covariance matrix of background errors with dimension  $n \times n$  ( $2n_r \times 2n_r$ ) is denoted  $\mathbf{B}$ . The background error standard deviations used in this study were  $\sigma_W = 0.707 \text{ g m}^{-3}$  and  $\sigma_{D_m} = 1 \text{ mm}$  (i.e., the variances are  $\sigma_W^2 = 0.5 \text{ g}^2 \text{ m}^{-6}$  and  $\sigma_{D_m}^2 = 1 \text{ mm}^2$ ). The relatively large background error gives more weight to observations, which is a reasonable assumption since this is an observation-based retrieval with no model background.

The background covariance assumes a Gaussian correlation model (Huang 2000):

$$b_{ij} = \sigma_b^2 \exp \left[ -\frac{1}{2} \left( \frac{r_{ij}}{r_L} \right)^2 \right], \quad (20)$$

where  $\sigma_b^2$  is the background error covariance ( $\sigma_W^2$  and  $\sigma_{D_m}^2$ ),  $r_{ij}$  is the distance between the  $i$ th and  $j$ th radar gates, and  $r_L$  is the spatial decorrelation length. In this study,  $r_L$  was set to 1000 m. An increase (decrease) in the spatial decorrelation length results in an increase (decrease) in smoothing of the final analysis. Therefore, the spatial influence of the observations is determined by the background error covariance matrix.

The covariance matrix of observations errors with dimension  $p \times p$  ( $3n_r \times 3n_r$ ) is denoted  $\mathbf{R}$ . The observation error standard deviations used in this study were  $\sigma_{Z_H} = 1 \text{ dB}$ ,  $\sigma_{Z_{DR}} = 0.2 \text{ dB}$ , and  $\sigma_{\Phi_{DP}} = 5^\circ$ . It is assumed that each observation's error is independent from the other observations. Therefore,  $\mathbf{R}$  is a diagonal matrix.

The observation operator  $H$  converts state variables in the analysis vector at iteration  $i$  to observations. These operators (11), (13), and (14) were derived in section 2.

The Jacobian  $\mathbf{H}$  contains the partial derivative of each observation operator with respect to the state variables. It has a dimension  $p \times n$  ( $3n_r \times 2n_r$ ). These can be calculated analytically by taking the partial derivatives of (11), (13), and (14) with respect to  $W$  and  $D_m$ :

$$\frac{\partial Z_H}{\partial W} = \frac{10}{W \ln(10)}, \quad (21)$$

$$\frac{\partial Z_H}{\partial D_m} = \frac{20(20.87 + 2 \times 46.04D_m - 3 \times 6.403D_m^2 + 4 \times 0.2248D_m^3)}{\ln(10)(-0.3078 + 20.87D_m + 46.04D_m^2 - 6.403D_m^3 + 0.2248D_m^4)}, \quad (22)$$

$$\frac{\partial Z_{DR}}{\partial W} = 0, \quad (23)$$

$$\frac{\partial Z_{DR}}{\partial D_m} = \frac{10(-0.1430 + 2 \times 0.3165D_m - 3 \times 0.06498D_m^2 + 4 \times 0.004163D_m^3)}{\ln(10)(1.019 - 0.1430D_m + 0.3165D_m^2 - 0.06498D_m^3 + 0.004163D_m^4)}, \quad (24)$$

$$\frac{\partial K_{DP}}{\partial W} = 0.009260 - 0.08699D_m + 0.1994D_m^2 - 0.02824D_m^3 + 0.001772D_m^4, \quad (25)$$

$$\frac{\partial K_{DP}}{\partial D_m} = W(-0.08699 + 2 \times 0.1994D_m - 3 \times 0.02824D_m^2 + 4 \times 0.001772D_m^3). \quad (26)$$

The partial derivative of  $\partial Z_{DR}/\partial W$  is zero because  $Z_{DR}$  only depends on  $D_m$ . For the partial derivatives  $\partial Z_H/\partial W$ ,  $\partial Z_H/\partial D_m$ , and  $\partial Z_{DR}/\partial D_m$ , the matrix is diagonal because they have gate-to-gate independence. For  $\partial \Phi_{DP}/\partial W$  and  $\partial \Phi_{DP}/\partial D_m$ , the matrix is a lower triangle because  $\Phi_{DP}$  is defined as 2 times the summation of  $K_{DP}$  up to a given gate:

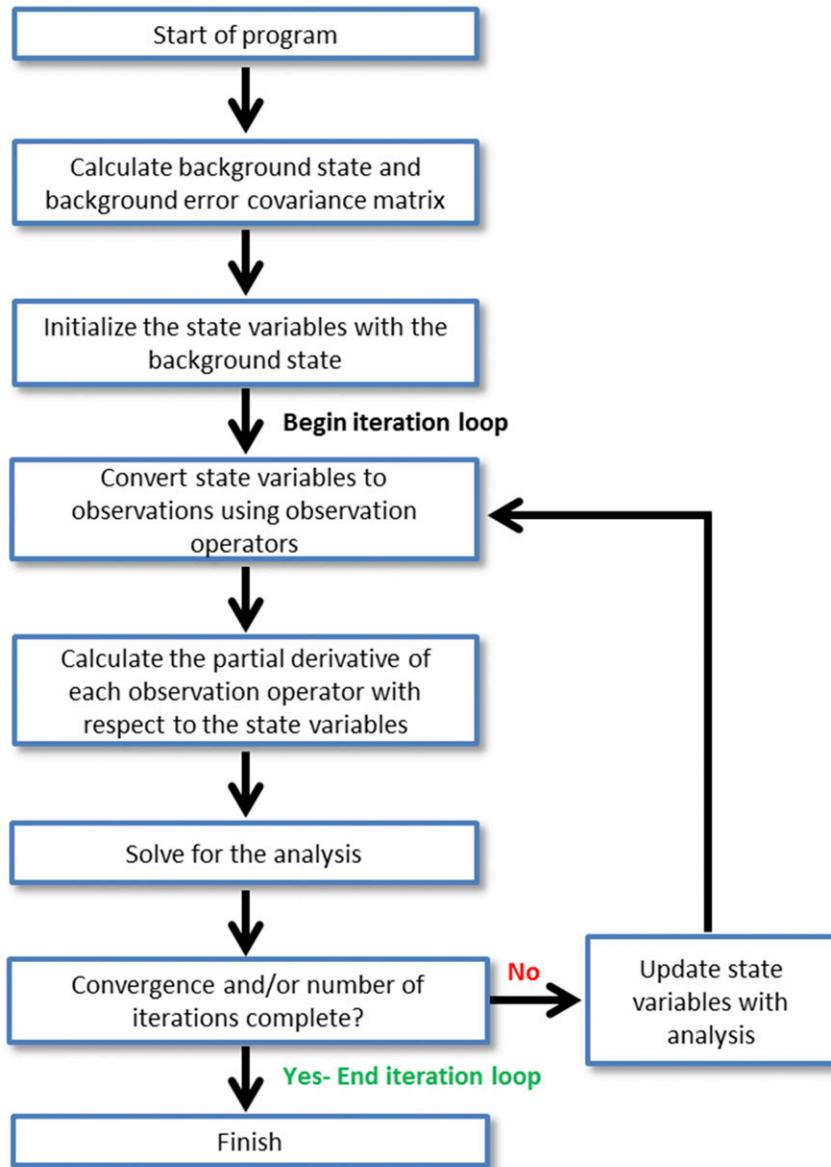


FIG. 2. Block diagram of the iteration procedure for the nonlinear variational retrieval for a single azimuth of radar data.

$$\Phi_{\text{DP}}(r_n) = 2 \sum_{i=1}^N K_{\text{DP}}(r_i) \Delta r_i. \quad (27)$$

The observation-based variational retrieval is complete when the analysis vector  $\mathbf{x}_r$  converges to a solution. A block diagram of the iteration procedure for the variational retrieval is shown in Fig. 2. In practice, the number of iterations can be fixed and/or a convergence test applied. If there are non-rain hydrometeors present (e.g., hail) anywhere in the azimuth, there may not be a solution or an unrealistic solution due to contamination within the azimuth (since the retrieval is solved azimuthally). This is because these parameterized polarimetric radar forward

observation operators were derived using the T-matrix calculations for rain.

To solve this for an entire plan position indicator (PPI) radar scan, each azimuth within a PPI scan is solved independently. A PPI scan is a radar scan where the elevation angle is constant and the azimuth angle varies. In other words, simply loop through all the azimuths to get a solution for an entire radar elevation scan.

#### 4. Testing the variational retrieval on DSD data

The variational retrieval and observation operators are tested by conducting experiments on 2DVD data

collected on 13 May 2005 (Zhang 2015). The 2DVD data are from a quasi-linear convective system (QLCS) that has an estimated storm motion from 16 to 18  $\text{m s}^{-1}$  or  $\sim 1 \text{ km min}^{-1}$ . The 2DVD data are collected for  $\sim 4.63 \text{ h}$ ; however, only the results for the first 60 min are shown because this is when the leading convective line of the QLCS passes over the disdrometer and when the most significant increase in  $\Phi_{\text{DP}}$  occurs (i.e., the most notable part of the dataset).

First, the intrinsic or true values for  $W$  and  $D_m$  are calculated using the 2DVD data. When calculating the intrinsic  $W$  and  $D_m$ , five-point averages of the raw  $W$  and  $D_m$  calculations are used to remove noise from the data. Second,  $Z_H$ ,  $Z_{\text{DR}}$ , and  $\Phi_{\text{DP}}$  are calculated using the observation operators [(11), (13), and (14)] and the relationship between  $\Phi_{\text{DP}}$  and  $K_{\text{DP}}$  [(27)]. These calculated polarimetric variables are assumed to be the true observations in these experiments. Third, it is assumed that the time series of 2DVD data represents a radial of radar data. This is assumed because the storm motion remained nearly steady state. Based on the storm motion, each minute of 2DVD data represents approximately a 1-km range gate of data. Thus, the entire 60 min of data represents a radial that is 60 km in length. Finally, in all of these experiments, a constant background is calculated using the empirical relationship given in (18) and (19). The mean of the empirically derived  $W$  and  $D_m$  are calculated using the entire radial of data ( $\sim 4.63 \text{ h}$ ). These mean values for  $W$  and  $D_m$  are used as the constant background.

#### a. OI versus nonlinear

In the first experiment, the OI solution (16) is compared to the nonlinear solution (17). Both of these solutions only include  $Z_H$  and  $Z_{\text{DR}}$  ( $\Phi_{\text{DP}}$  is not included). The relative advantage of the nonlinear solution over OI is demonstrated in Fig. 3. For all analyses, the nonlinear solution is closer to the truth. For example, the peak values of  $W$  are 0.65 and  $1.58 \text{ g m}^{-3}$  for the OI and nonlinear solutions at 20 km, respectively (Fig. 3a). The nonlinear solution is closer to the true maximum value of  $1.94 \text{ g m}^{-3}$ .

For  $Z_H$ , there is generally an underestimation of values in the OI solution when compared to the nonlinear solution and truth (Fig. 3b). For example, the peak values are  $\sim 45$  and  $\sim 50 \text{ dBZ}$  for the OI and nonlinear solutions at 29 km, respectively. The nonlinear solution is the same as the truth in this case.  $D_m$  and  $Z_{\text{DR}}$  are similar for both the OI and nonlinear solutions with an exception of the peak values (Figs. 3c,d). There a slight underestimation on the OI solution at peak values. For example, the true value at 12 km is 3.05 dB for  $Z_{\text{DR}}$  and 3.00 mm for  $D_m$ . This compares to 2.85 dB and 2.80 mm

for the OI solution and 3.00 dB and 2.94 mm for the nonlinear solution.

The advantage of the nonlinear solution over OI is also shown by the final value of  $\Phi_{\text{DP}}$ . Since  $\Phi_{\text{DP}}$  is integrated in the along-radial direction, the value at the final gate provides insight on the performance across the entire radial of data. In this experiment, the final value of  $\Phi_{\text{DP}}$  for the OI solution at  $\sim 15^\circ$  is nearly half when compared to the nonlinear solution at  $\sim 29^\circ$  (Fig. 3e). The nonlinear solution is very close to the true value, which is  $30^\circ$ . This seems reasonable and intuitive based on the significant underestimation of  $W$  across the entire radial with the OI solution.

#### b. Nonlinear without $\Phi_{\text{DP}}$ versus nonlinear with $\Phi_{\text{DP}}$

In the second experiment, the nonlinear solution without  $\Phi_{\text{DP}}$  is compared to that with  $\Phi_{\text{DP}}$ . The advantage of including  $\Phi_{\text{DP}}$  in the variational retrieval is shown in Fig. 4. The most evident advantage is seen in  $W$ , especially with larger values ( $> 1.5 \text{ g m}^{-3}$ ) (Fig. 4a). The peak values of  $W$  are 1.58 and  $1.75 \text{ g m}^{-3}$  for nonlinear solution without and with  $\Phi_{\text{DP}}$ , respectively. The nonlinear solution with  $\Phi_{\text{DP}}$  is closer to the true maximum value of  $1.94 \text{ g m}^{-3}$ .

For  $Z_H$ ,  $D_m$ , and  $Z_{\text{DR}}$ , the benefits of including  $\Phi_{\text{DP}}$  are negligible (Figs. 4b,c,d). The final value of  $\Phi_{\text{DP}}$  for the nonlinear solution without  $\Phi_{\text{DP}}$  is  $\sim 29^\circ$ , which is close to the nonlinear solution with  $\Phi_{\text{DP}}$  and the truth at  $\sim 30^\circ$  (Fig. 4e). When compared to the previous experiment, the final value of  $\Phi_{\text{DP}}$  does not have a significant difference.

Even though most of the analyses show little or no benefit, it is important to note that there is still evidence of some benefit of adding  $\Phi_{\text{DP}}$  in addition to  $Z_H$  and  $Z_{\text{DR}}$  as shown with the increase in the peak  $W$  value. This improves results for  $W$  that is closer to the true value. In other words, the inclusion of  $\Phi_{\text{DP}}$  is important for the retrieval of larger  $W$  values.

#### c. Simulated observations

In the third experiment, the nonlinear solution with  $\Phi_{\text{DP}}$  is compared to the nonlinear solution with  $\Phi_{\text{DP}}$  with random error added to the observations. Adding random error may result in worse retrievals, which is an important consideration since true radar observations include random error. The random errors included are  $\pm 1 \text{ dBZ}$  for  $Z_H$ ,  $\pm 0.2 \text{ dB}$  for  $Z_{\text{DR}}$ , and  $\pm 5^\circ$  for  $\Phi_{\text{DP}}$ . These values were chosen because they are used in the covariance matrix of observations errors, as described in previous section.

As shown in Fig. 5, this experiment demonstrates that the nonlinear variational retrieval still provides a reasonable solution even with random error in the observations. This is especially true when looking at the solution as a whole. As previously mentioned, the

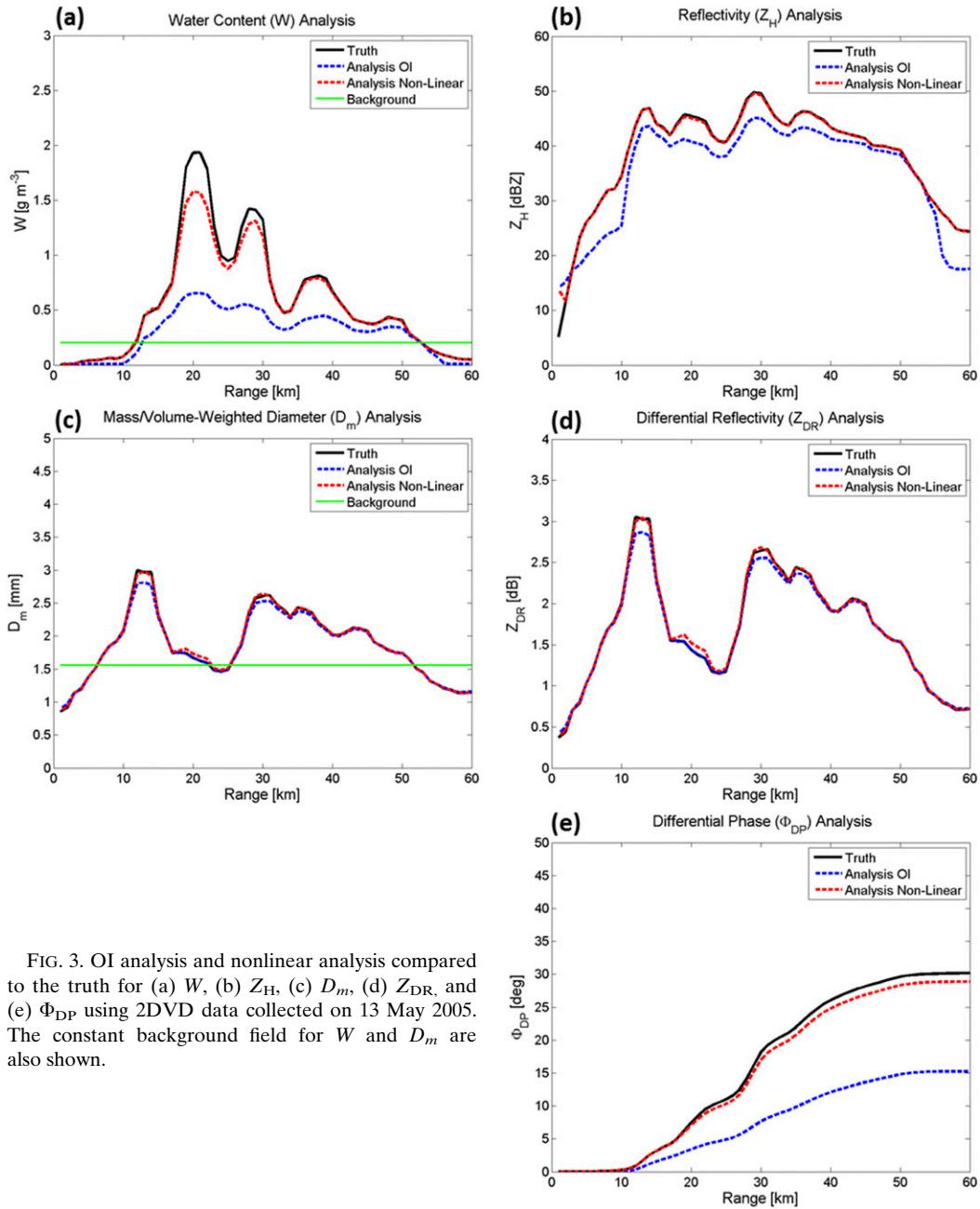


FIG. 3. OI analysis and nonlinear analysis compared to the truth for (a)  $W$ , (b)  $Z_H$ , (c)  $D_m$ , (d)  $Z_{DR}$ , and (e)  $\Phi_{DP}$  using 2DVD data collected on 13 May 2005. The constant background field for  $W$  and  $D_m$  are also shown.

best way to assess this is by looking at the final  $\Phi_{DP}$  value because  $\Phi_{DP}$  is an integrated quantity. The final value of  $\Phi_{DP}$  with random error is the same as truth ( $\sim 30^\circ$ ) (Fig. 5e). Therefore, on average, the solution is still reasonable.

**5. Applying the variational retrieval to radar data**

In the previous section, the variational retrieval and observation operators were successfully tested on 2DVD data and simulated observations. The next step is to apply

them to real radar data. For this study, the variational retrieval is applied to the PRD collected from the S-band Norman, Oklahoma (KOUN), Weather Surveillance Radar-1988 Doppler (WSR-88D) on 15 June 2011. The radar data are obtained from the National Centers for Environmental Information (NOAA/NCEI 2011). The radar data from this event captures observations of a downburst and its parent thunderstorm. Only the PRD at a range up to  $\sim 32$  km is used for the variational retrieval because this is a sufficient range to capture the storm of interest. More information about the downburst

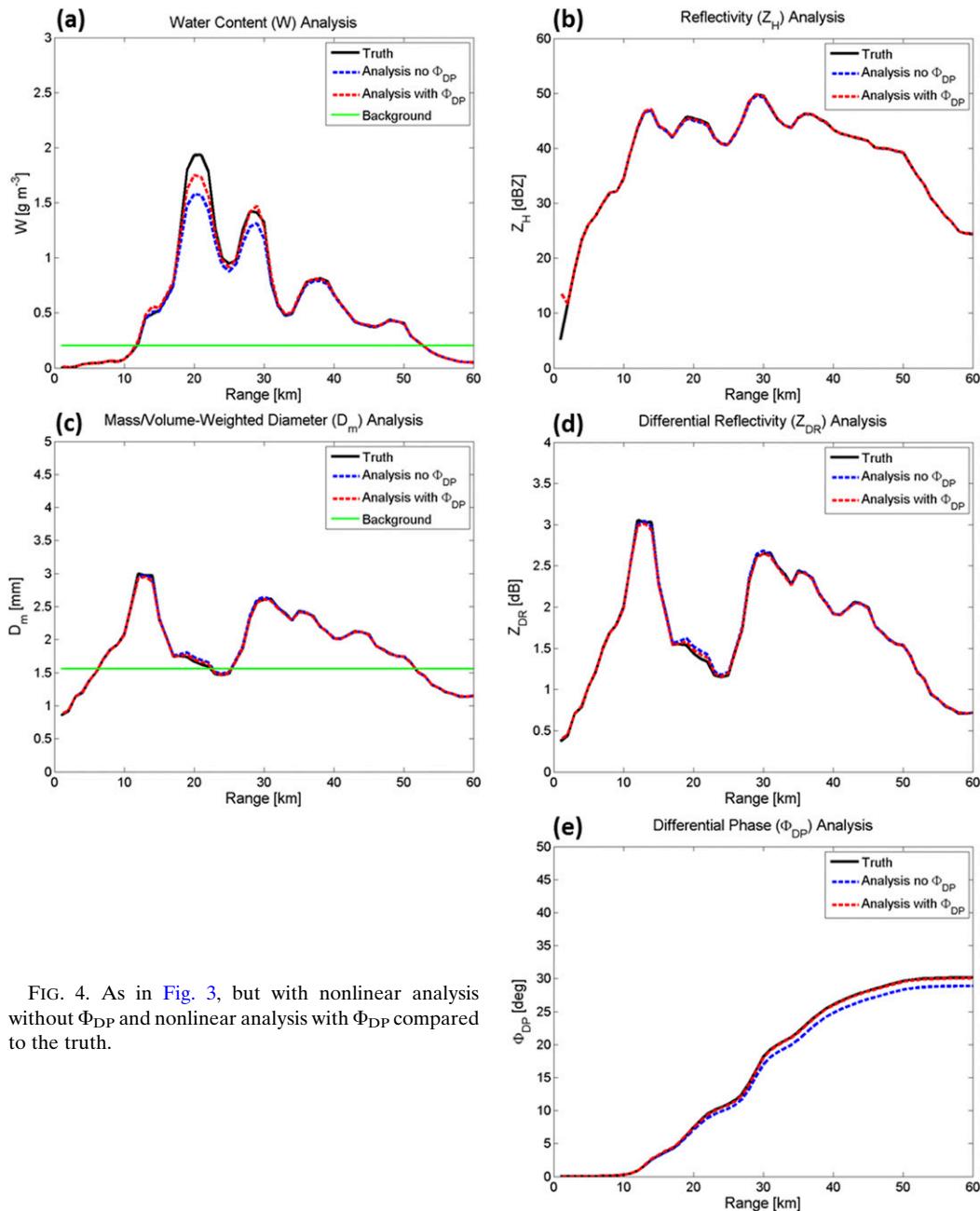


FIG. 4. As in Fig. 3, but with nonlinear analysis without  $\Phi_{DP}$  and nonlinear analysis with  $\Phi_{DP}$  compared to the truth.

event can be found in studies by Kuster et al. (2016) and Mahale et al. (2016).

#### a. Single-azimuth experiment

The variational retrieval is tested on a single azimuth of radar data before applying it to an entire PPI scan. As previously mentioned, the variational retrieval is solved azimuthally (1D) to coincide with  $\Phi_{DP}$  observations. As with the 2DVD data, some assumptions and preprocessing to the data are applied to this experiment.

A constant background for  $W$  and  $D_m$  are calculated using the azimuthal mean of empirical relationship given in (18) and (19). Also as with 2DVD data, quality-control of the radar data is accomplished through a median filter to reduce the noise in the data. In addition, some constraints are placed on the radar data. The  $Z_H$  minimum is set to 10 dBZ,  $Z_{DR}$  is limited to 0.1 to 6 dB, and only positive values of  $\Phi_{DP}$  are allowed. These constraints are based on the bounds of the observation operators as shown in Fig. 1. Finally, a simple attenuation

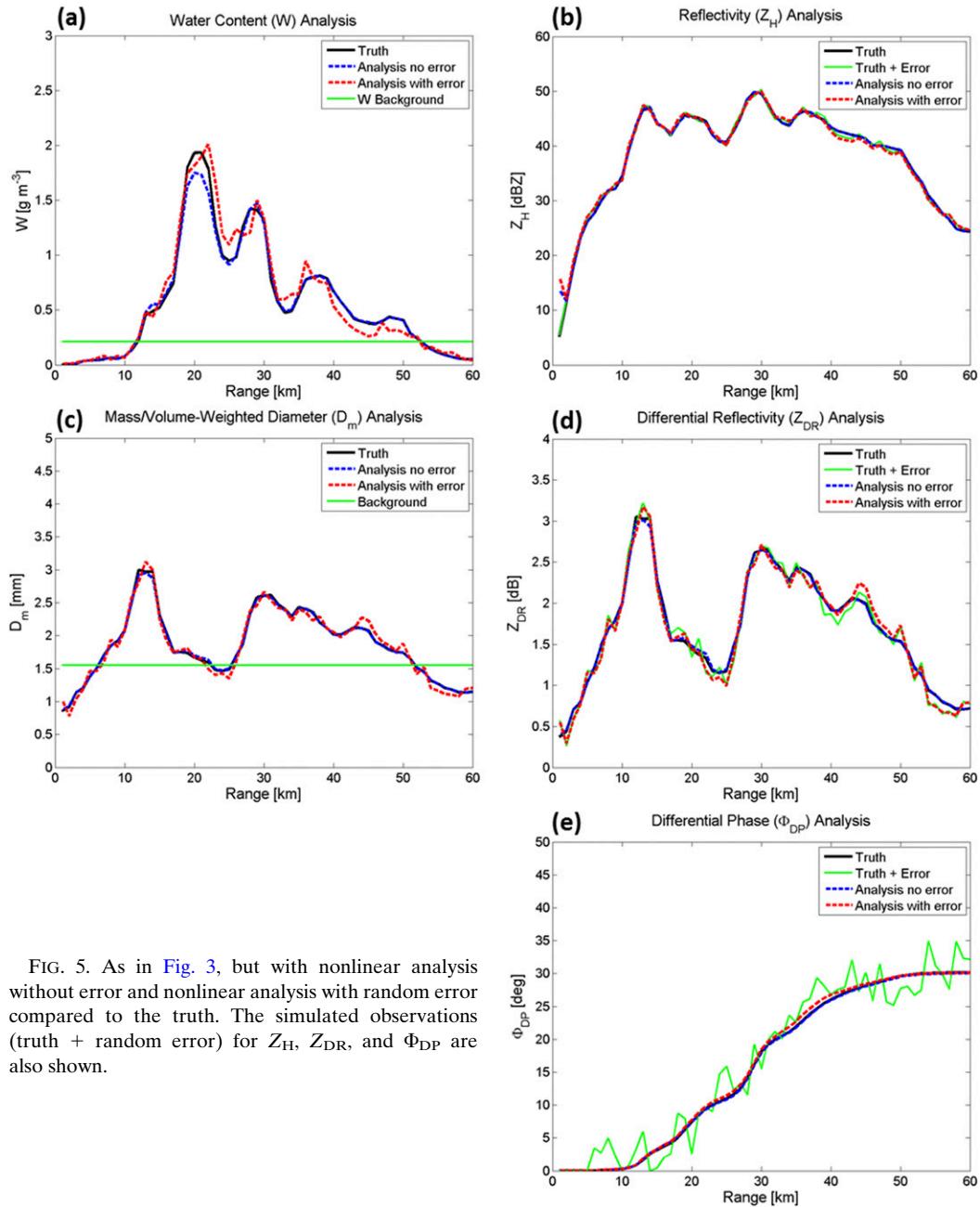


FIG. 5. As in Fig. 3, but with nonlinear analysis without error and nonlinear analysis with random error compared to the truth. The simulated observations (truth + random error) for  $Z_H$ ,  $Z_{DR}$ , and  $\Phi_{DP}$  are also shown.

correction is also implemented to the radar data using the differential phase (DP) attenuation correction procedure (Bringi et al. 1990) with coefficients derived in Zhang (2016).

In this experiment, PRD from an azimuth ( $\sim 3.5^\circ$ ) with moderate  $Z_H$  ( $\sim 30\text{--}50$  dBZ) and high  $\rho_{hv}$  ( $>0.98$ ) values is chosen. Based on these observations, the assumption is that only raindrops associated moderate rainfall are present in this azimuth with no hail contamination. This makes it an ideal azimuth to test the

variational retrieval. Once  $W$  and  $D_m$  are solved using the variational retrieval, the observation operators are used to calculate  $Z_H$ ,  $Z_{DR}$ , and  $\Phi_{DP}$  analyses as in the 2DVD experiments.

The  $W$ ,  $D_m$ ,  $Z_H$ ,  $Z_{DR}$ , and  $\Phi_{DP}$  analyses for the single azimuth are shown in Fig. 6. For comparison, the empirical relationships (18) and (19) are used to calculate  $W$  and  $D_m$ . The empirical relationships only depend on  $Z_H$  and  $Z_{DR}$  and do not account for  $\Phi_{DP}$ . Even so, the empirical relationships can be used to qualitatively assess the

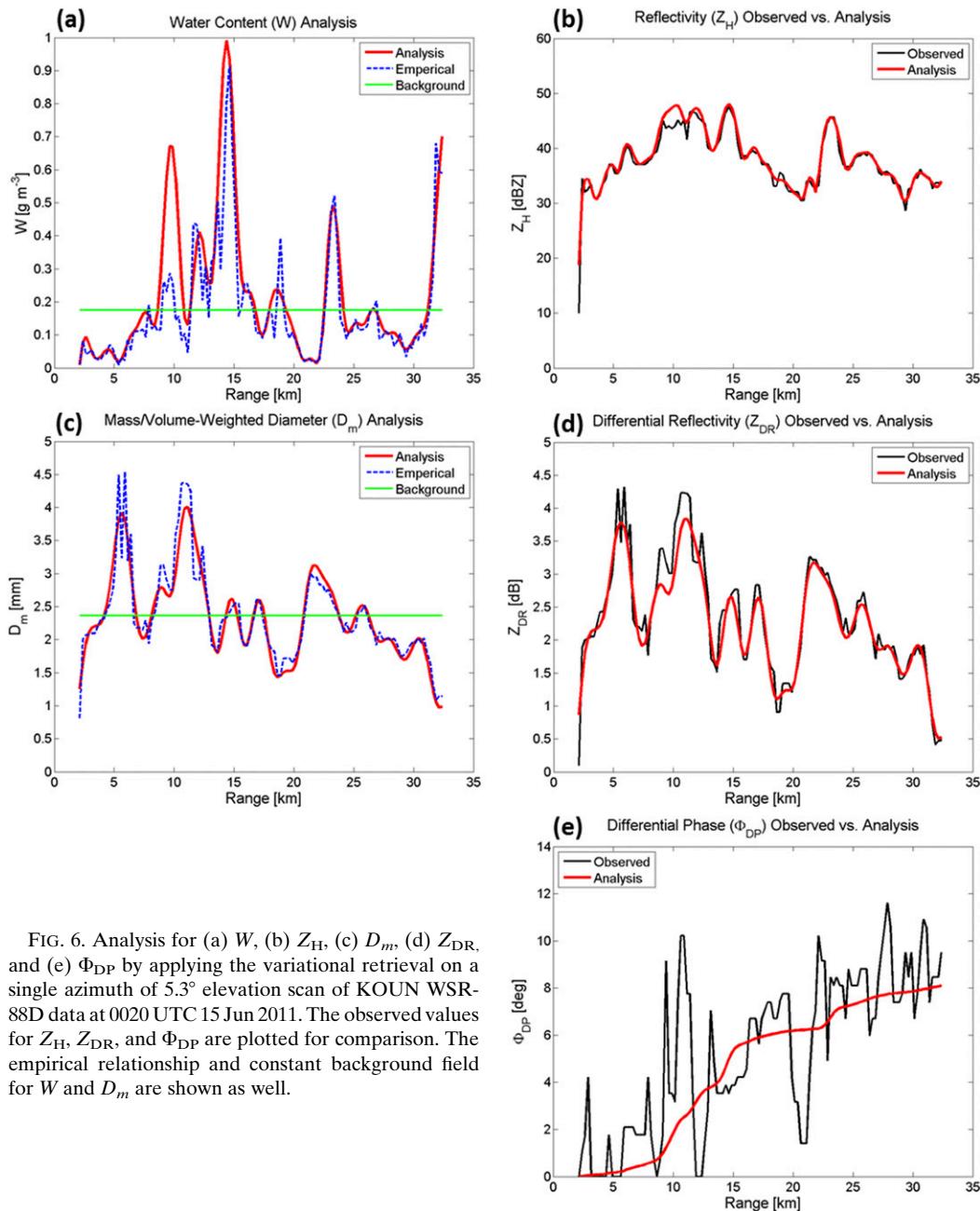


FIG. 6. Analysis for (a)  $W$ , (b)  $Z_H$ , (c)  $D_m$ , (d)  $Z_{DR}$ , and (e)  $\Phi_{DP}$  by applying the variational retrieval on a single azimuth of  $5.3^\circ$  elevation scan of KOUN WSR-88D data at 0020 UTC 15 Jun 2011. The observed values for  $Z_H$ ,  $Z_{DR}$ , and  $\Phi_{DP}$  are plotted for comparison. The empirical relationship and constant background field for  $W$  and  $D_m$  are shown as well.

results of the variational retrieval. The observed values for  $Z_H$ ,  $Z_{DR}$ , and  $\Phi_{DP}$  are shown as a comparison as well.

For  $W$ , the variational retrieval is smoother than the empirical relationship (Fig. 6a). This is reasonable because the variational retrieval takes account for the spatial covariance, which is a function of the spatial decorrelation length. The variational retrieval would be less smooth (smoother) with a smaller (larger) spatial decorrelation length. The local maxima and minima for  $W$  are at approximately the same range;

however, the magnitudes of these maxima and minima are sometimes different. For example, at  $\sim 9$ – $10$ -km range, the maximum magnitude of the difference between the two  $W$  values is  $0.43 \text{ g m}^{-3}$ .

This range is also where there is a noticeable difference between the  $Z_H$  analysis and observations (Fig. 6b). At the same range, the maximum magnitude of the difference between  $Z_H$  analysis and observations is 3.8 dB. The  $\Phi_{DP}$  observations are essentially increasing  $W$  values beyond simply what the  $Z_H$  values

would suggest in this range, which results in larger  $Z_H$  values in the analysis to keep it consistent with the  $Z_H$  observation operator. If  $\sigma_{\Phi_{DP}}$  was increased (decreased), the  $W$  values in this range would decrease (increase) because of decreasing (increasing) influence of  $\Phi_{DP}$  with more (less) weight given to  $Z_H$ . Outside of this range, the  $Z_H$  analysis is fairly close to the  $Z_H$  observations (albeit smoother).

For the  $D_m$  analysis (Fig. 6c), the variational retrieval is smoother than the empirical relationship as well and the peak values are less than the empirical relationship at a range of  $\sim 6$  and  $\sim 11$  km. This occurs because the observation operators, and more specifically the T-matrix calculations, have a smaller dynamic range for  $D_m$ . Consequently, the reduction of  $D_m$  also results in a coincident reduction in  $Z_{DR}$  as well (Fig. 6d). Outside these areas, the  $Z_{DR}$  analysis is fairly close to the  $Z_{DR}$  observations.

When comparing  $W$  with  $\Phi_{DP}$  at this range (Figs. 6a,e), it is evident that the  $\Phi_{DP}$  has a positive impact in modulating  $W$ . The  $\Phi_{DP}$  analysis has a relatively steep slope in this range (i.e., relatively large  $K_{DP}$ ), which is why  $W$  is greater than in the empirical relationship that does not have the contribution from  $\Phi_{DP}$ .

It is also noteworthy that the  $\Phi_{DP}$  analysis can only monotonically increase when using the variational retrieval (Fig. 6e). This is because only positive  $K_{DP}$  values are allowed in the observation operator. As shown,  $\Phi_{DP}$  observations are noisy and chaotic when compared to  $Z_H$  and  $Z_{DR}$  observations, which is why  $\sigma_{\Phi_{DP}}$  is relatively large when compared to  $\sigma_{Z_H}$  and  $\sigma_{Z_{DR}}$ . The variational retrieval provides a method to calculate a smoother, monotonically increasing  $\Phi_{DP}$  while accounting for the  $Z_H$  and  $Z_{DR}$  observations. This is the characteristic that would be expected for  $\Phi_{DP}$  in a pure-rain azimuth with no hail contamination at S band.

In addition, because the  $\Phi_{DP}$  analysis can only monotonically increase with range,  $\Phi_{DP}$  observations serve as a constraint for the variational retrieval. This is because the final  $\Phi_{DP}$  analysis value is dependent on the total sum of  $W$  across the entire azimuth. The other two observations,  $Z_H$  and  $Z_{DR}$ , are independent from gate to gate when attenuation and differential attenuation can be ignored.

Overall, the single-azimuth experiment on an area of moderate rainfall was successful. The locations of local maxima and minima for  $W$ ,  $D_m$ ,  $Z_H$ , and  $Z_{DR}$  are similar to either the empirical formula or the observations with some differences in magnitude. The similarity to the empirical formula provides a proof of concept that the variational retrieval yields a reasonable solution. The subtle differences can be accounted for mostly with the addition of  $\Phi_{DP}$  observations (in the case of  $W$  and  $Z_H$ ) and the limitation of T-matrix

calculations used in the observation operators (in the case of  $D_m$  and  $Z_{DR}$ ). Finally, these results demonstrate that the variational retrieval has utility in calculating a smoother, monotonically increasing (more physically representing)  $\Phi_{DP}$ .

### b. PPI experiment

After the single-azimuth experiment, the variational retrieval was applied on the entire PPI of radar data by looping through all 360 azimuths. As with the single-azimuth experiment, once  $D_m$  and  $W$  are obtained using the variational retrieval, the observation operators are used to calculate  $Z_H$  and  $Z_{DR}$  analyses to assess the observation operators in context to the forward problem. Instead of calculating  $\Phi_{DP}$ , the  $K_{DP}$  analysis is shown to demonstrate it can be estimated using this method.  $K_{DP}$  is operationally useful because it is better correlated to rainfall rates due to it being unaffected by calibration error and attenuation as well as being more linearly related to rainfall rate, which gives it utility for quantitative precipitation estimation (QPE) (Bringi and Chandrasekar 2001). The  $W$  and  $D_m$  analyses are shown in Fig. 7 and the  $Z_H$  and  $Z_{DR}$  observations and analyses are shown in Fig. 8. The differences between the observations and analyses are also shown (Figs. 8e,f). The  $K_{DP}$  analysis is shown in Fig. 9.

The most evident feature in the analyses is the noisy results across the western to northwestern sector of the PPI. Parts of the analyses in this area are unrealistic and the  $Z_H$  analysis differs significantly from the observations. There are even some azimuths that are completely void of data in the analyses. Further investigation was done on why the variational retrieval failed in this sector and will be discussed later in this section. In addition to this sector, there are a few other azimuths that are void of data. These azimuths do not have a solution likely due to low  $Z_H$  values and/or the presence of non-meteorological scatterers (i.e., biological or ground clutter) in the azimuth.

Outside of these areas, the  $Z_H$  and  $Z_{DR}$  analyses are similar to the observations (Fig. 8). The local maxima and minima are near the same locations (with an occasional variation in the magnitude), which results in the same precipitation structure. Some variability in magnitude is expected because an optimal analysis that utilizes all the information from  $Z_H$ ,  $Z_{DR}$ , and  $\Phi_{DP}$ . Also, some smoothing occurs because the spatial covariance is taken into account. One noticeable difference is that the peak magnitude of  $Z_{DR}$  is tempered when compared to observations. As noted in the single-azimuth experiment, this is because of the dynamic range of  $D_m$  that can be calculated through the variational retrieval.

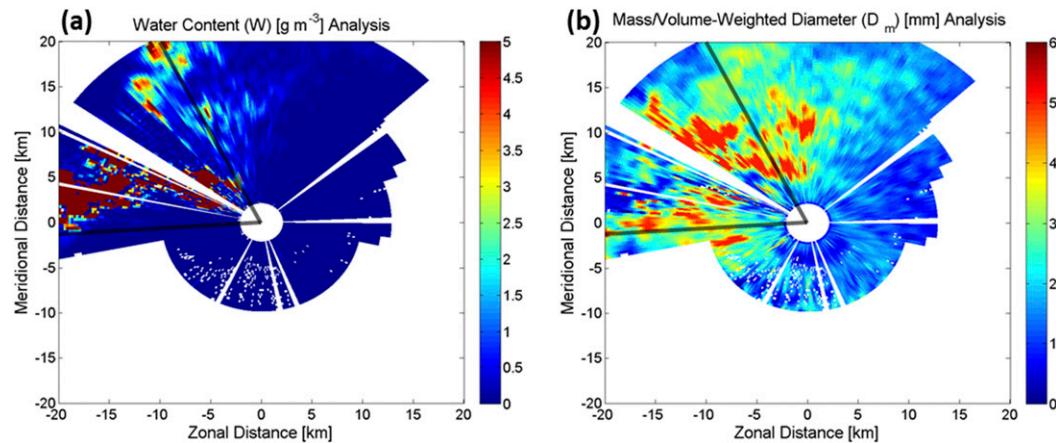


FIG. 7. Analysis for (a)  $W$  and (b)  $D_m$  by applying the variational retrieval using a  $5.3^\circ$  elevation scan of KOUN WSR-88D data at 0020 UTC 15 Jun 2011. The boundaries of the sector with hail contamination are marked by the black lines.

There are also slight discontinuities from azimuth to azimuth. This is likely an artifact from changes in the errors of  $\Phi_{DP}$  observations from azimuth to azimuth and the subsequent effect on the final  $\Phi_{DP}$  analysis value. In this study, the  $\Phi_{DP}$  error distribution is assumed constant across the entire PPI, which may not be the case. As previously noted, the final  $\Phi_{DP}$  analysis value is dependent on the total sum of  $W$  across the entire azimuth. Slight deviations in the final  $\Phi_{DP}$  analysis value will result in a slight change for the retrieved state variables across the entire azimuth.

The  $K_{DP}$  analysis is an optimal analysis that utilizes information from  $Z_H$  and  $Z_{DR}$  (in addition to  $\Phi_{DP}$ ) (Fig. 9). As previously noted,  $\Phi_{DP}$  can only monotonically increase when using the variational retrieval. This is a constraint that comes out of the observation operator. Therefore, negative  $K_{DP}$  values cannot exist. This is reasonable for areas of pure rain. Like  $Z_H$  and  $Z_{DR}$ ,  $K_{DP}$  is relatively smooth from the variational retrieval.

Applying a simplified hydrometeor classification algorithm (HCA) from Mahale et al. (2014) (based upon Park et al. 2009) provided more insight on why the variational retrieval failed on the western to northwestern sector of the PPI. The HCA discriminates between 10 classes of radar echo: 1) ground clutter and anomalous propagation (GC/AP), 2) biological scatterers (BS), 3) dry aggregated snow (DS), 4) wet snow (WS), 5) crystals (CR), 6) graupel (GR), 7) big drops (BD), 8) light and moderate rain (RA), 9) heavy rain (HR), and 10) a mixture of rain and hail (RH). As shown in Fig. 10, these radials had radar gates that are classified as RH, which implies there was potentially hail contamination on these radials. This classification seems reasonable based on the observed  $Z_H$  and  $Z_{DR}$  values. Relatively high  $Z_H$  with near-zero or negative  $Z_{DR}$  has been observed as a

signature associated with hail for S-band radars (e.g., Bringi et al. 1986). In addition, as previously mentioned, the HCA also indicates GC/AP and/or BS in some of the other azimuths where no solution was found.

As noted, if there is hail contamination (or any non-rain hydrometers) present anywhere in the azimuth, there may not be a solution or an unrealistic solution due to contamination. This is because the observation operators were derived using the T-matrix calculations for pure rain. On Figs. 7–9, the sector associated with hail contamination is highlighted using the results from the HCA. For operational purposes in the detection of hazardous weather, the failure of the variational retrieval could actually be a tool used to determine if hail is present in a thunderstorm. This would be especially true if the variational retrieval is conducted locally over a set number of simultaneous gates rather than an entire azimuth.

## 6. Summary and conclusions

In this study, a proposed method for the observation-based variational retrieval of the physical parameters of  $W$  and  $D_m$  from S-band PRD was developed through the use of parameterized polarimetric radar observation operators and a nonlinear, iterative method.

The observation operators were derived for  $Z_H$ ,  $Z_{DR}$ , and  $K_{DP}$  using the T-matrix method and an exponential distribution model, which link PRD to  $D_m$  and  $W$ . These simple observation operators are advantageous because they include relevant physics, are numerically efficient, and have easily calculated first derivatives. In addition, these parameters can be easily assimilated to NWP because are directly related to the prognostic physics variables of  $N_r$  and  $q_r$ .

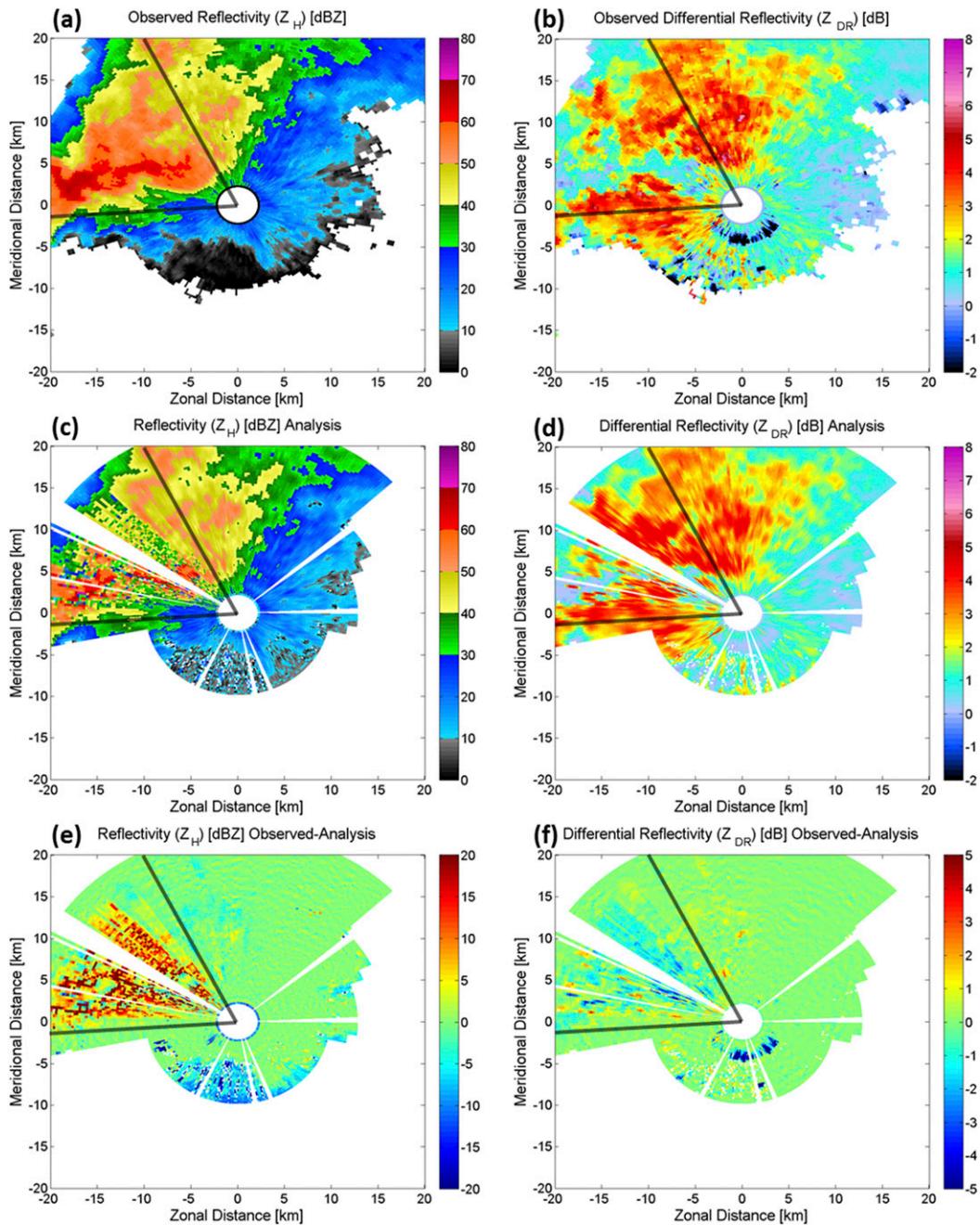


FIG. 8. The (a)  $Z_H$  observations, (b)  $Z_{DR}$  observations, (c)  $Z_H$  analysis, (d)  $Z_{DR}$  analysis, (e)  $Z_H$  observations and analysis difference, and (f)  $Z_{DR}$  observations and analysis difference using a  $5.3^\circ$  elevation scan of KOUN WSR-88D data at 0020 UTC 15 Jun 2011. The analyses are from applying the variational retrieval.

The use of an exponential distribution results in an overdetermined problem, which is advantageous in this retrieval because  $\Phi_{DP}$  is used as a constraint along the azimuth (i.e., it can only monotonically increase), the PRD are not all independent (e.g.,  $Z_{DR}$  has dependence on  $Z_H$ ), and there are measurement and model errors. For example, this would make the retrieval less sensitive

to  $Z_{DR}$  bias because  $Z_H$ ,  $Z_{DR}$ , and  $K_{DP}$  are all a function of  $D_m$  (though some error will be added to the retrieval due to the assumption that observations are unbiased). The disadvantage of an exponential distribution is that it may not perform as well in areas where there is concavity present in the DSD. Nevertheless, recent work has shown the two parameter DSD model has an advantage for

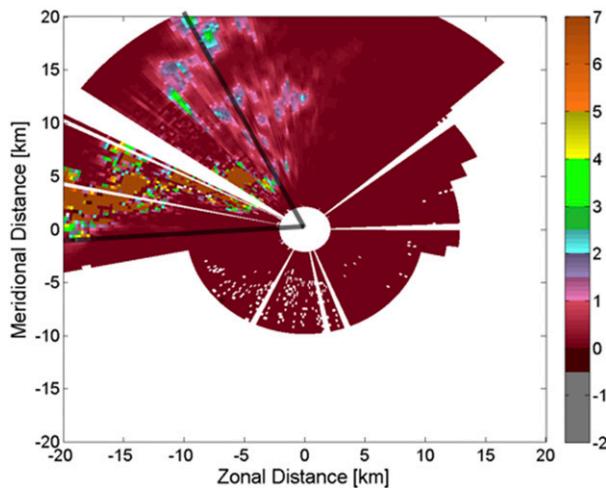


FIG. 9. Analysis for  $K_{DP}$  from the variational retrieval using a  $5.3^\circ$  elevation scan of KOUN WSR-88D data at 0020 UTC 15 Jun 2011. The boundaries of the sector with hail contamination are marked by the black lines.

retrieving DSDs from PRD when compared to the three-parameter gamma DSD model (Huang et al. 2019).

Using these observation operators, the Gauss–Newton iterative method for moderately nonlinear problems was successively used in the variational retrieval. Experiments on calculated  $W$  and  $D_m$  from 2DVD data demonstrated the advantages of using the nonlinear solution when compared to the OI solution and including  $\Phi_{DP}$  observations in addition to  $Z_H$  and  $Z_{DR}$ . The  $\Phi_{DP}$  observations may have only provided a slight benefit because they do not increase substantially at S band (especially for light to moderate rainfall). It was also shown even with the inclusion of random error into the 2DVD data, a reasonable solution can still be found. These advantages were discussed in context of the both the retrieval of  $W$  and  $D_m$  and the estimation of polarimetric radar variables to test the observation operators both inverse and forward, respectively. It can be concluded from these experiments that accounting for nonlinearity and including  $\Phi_{DP}$  observations is beneficial for the retrieval  $W$  and  $D_m$ . The nonlinearity is noteworthy because even though most problems are not truly linear, the data assimilation community often assumes linearity.

When applied to real radar data, the variational method yields promising results for the retrieved  $W$  and  $D_m$  and derived  $K_{DP}$  within areas of pure rain. Since the observation operators were derived with the assumption of pure rain, the retrieval failed where there was hail contamination. In this case, the derived  $K_{DP}$  from  $W$  and  $D_m$  can be considered a related parameter that describes the rain microphysics.

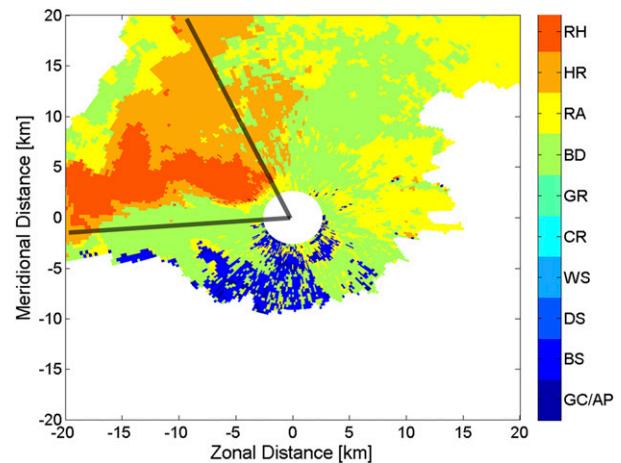


FIG. 10. HCA applied to  $5.3^\circ$  elevation scan of KOUN WSR-88D data at 0020 UTC 15 Jun 2011. The classifications used in HCA are defined in section 5b. The boundaries of the sector with hail contamination are marked by the black lines.

Overall, the successful variational retrieval of  $W$  and  $D_m$  from the 2DVD and the real radar data from two separate events demonstrate the utility of the proposed method. This method develops a framework that utilizes proven techniques from the data assimilation community. While the inclusion of a background state may not be a significant advantage for 1D (azimuth) retrieval, it can be important for 2D or 3D retrieval problems, which can become underdetermined if the background is not included because of the data distribution/coverage and spacing.

In the future, additional studies can be conducted on simulated and real data using this proposed retrieval method. This may include controlled experiments where we know the true quality of the measurements and their error. Other future research includes deriving the observation operators for other radar frequencies (e.g., X and C bands) and incorporating observation operators for other hydrometer types such as hail. The inclusion of other hydrometeors is important for operational use, especially when retrieving information above the melting layer. For X and C bands, attenuation would have to definitely be accounted for in the  $Z_H$  and  $Z_{DR}$  observation operators (or applied prior to the retrieval). Finally, a goal is to utilize the variational retrieval for data assimilation in NWP to improve model forecasts.

*Acknowledgments.* This work was partially supported by NOAA Grant NOAA-OAR-OWAQ-2017-2005122 and NOAA Grant NA16OAR4320115. Jidong Gao was partially supported by NSF AGS-1341878. The authors would like to acknowledge the National Severe Storms

Laboratory for collecting the KOUN WSR-88D data. The authors thank the three anonymous reviewers who provided constructive feedback on the original version of this paper. This research represents a portion of Vivek Mahale's Ph.D. dissertation at the University of Oklahoma.

## REFERENCES

- Bringi, V. N., and V. Chandrasekar, 2001: *Polarimetric Doppler Weather Radar: Principles and Applications*. Cambridge University Press, 636 pp.
- , J. Vivekanandan, and J. D. Tuttle, 1986: Multiparameter radar measurements in Colorado convective storms. Part II: Hail detection studies. *J. Atmos. Sci.*, **43**, 2564–2577, [https://doi.org/10.1175/1520-0469\(1986\)043<2564:MRMICC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1986)043<2564:MRMICC>2.0.CO;2).
- , V. Chandrasekar, N. Balakrishnan, and D. S. Zrnic, 1990: An examination of propagation effects in rainfall on polarimetric variables at microwave frequencies. *J. Atmos. Oceanic Technol.*, **7**, 829–840, [https://doi.org/10.1175/1520-0426\(1990\)007<0829:AEPEI>2.0.CO;2](https://doi.org/10.1175/1520-0426(1990)007<0829:AEPEI>2.0.CO;2).
- Cao, Q., G. Zhang, E. Brandes, and T. Schuur, 2010: Polarimetric radar rain estimation through retrieval of drop size distribution using a Bayesian approach. *J. Appl. Meteor. Climatol.*, **49**, 973–990, <https://doi.org/10.1175/2009JAMC2227.1>.
- , —, and M. Xue, 2013: A variational approach for retrieving raindrop size distribution from polarimetric radar measurements in the presence of attenuation. *J. Appl. Meteor. Climatol.*, **52**, 169–185, <https://doi.org/10.1175/JAMC-D-12-0101.1>.
- Eliassen, A., 1954: Provisional report on calculation of spatial covariance and autocorrelation of the pressure field. Rep. 5, Videnskaps-Akademiet, Institut for Vaeroch Klimaforskning, 12 pp.
- Gandin, L., 1963: *Objective Analysis of Meteorological Fields*. Gidromet, 285 pp.
- Gao, J., M. Xue, A. Shapiro, and K. K. Droegemeier, 1999: A variational method for the analysis of three-dimensional wind fields from two Doppler radars. *Mon. Wea. Rev.*, **127**, 2128–2142, [https://doi.org/10.1175/1520-0493\(1999\)127<2128:AVMFTA>2.0.CO;2](https://doi.org/10.1175/1520-0493(1999)127<2128:AVMFTA>2.0.CO;2).
- , —, —, Q. Xu, and K. K. Droegemeier, 2001: Three-dimensional simple adjoint velocity retrievals from single-Doppler radar. *J. Atmos. Oceanic Technol.*, **18**, 26–38, [https://doi.org/10.1175/1520-0426\(2001\)018<0026:TDSAVR>2.0.CO;2](https://doi.org/10.1175/1520-0426(2001)018<0026:TDSAVR>2.0.CO;2).
- Hogan, R. J., 2007: A variational scheme for retrieving rainfall rate and hail reflectivity fraction from polarization radar. *J. Appl. Meteor. Climatol.*, **46**, 1544–1564, <https://doi.org/10.1175/JAM2550.1>.
- Huang, H., G. Zhang, K. Zhao, S. Liu, L. Wen, G. Chen, and Z. Yang, 2019: Uncertainty in retrieving raindrop size distribution from polarimetric radar measurements. *J. Atmos. Oceanic Technol.*, **36**, 585–605, <https://doi.org/10.1175/JTECH-D-18-0107.1>.
- Huang, X. Y., 2000: Variational analysis using spatial filters. *Mon. Wea. Rev.*, **128**, 2588–2600, [https://doi.org/10.1175/1520-0493\(2000\)128<2588:VAUSF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<2588:VAUSF>2.0.CO;2).
- Ide, K., P. Courtier, M. Ghil, and A. C. Lorenc, 1997: Unified notation for data assimilation: Operational, sequential and variational. *J. Meteor. Soc. Japan*, **75**, 181–189, [https://doi.org/10.2151/jmsj1965.75.1B\\_181](https://doi.org/10.2151/jmsj1965.75.1B_181).
- Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation, and Predictability*. Cambridge University Press, 341 pp.
- Kruger, A., and W. F. Krajewski, 2002: Two-dimensional video disdrometer: A description. *J. Atmos. Oceanic Technol.*, **19**, 602–617, [https://doi.org/10.1175/1520-0426\(2002\)019<0602:TDVDAD>2.0.CO;2](https://doi.org/10.1175/1520-0426(2002)019<0602:TDVDAD>2.0.CO;2).
- Kuster, C. M., P. L. Heinselman, and T. J. Schuur, 2016: Rapid update radar observations of downbursts occurring within an intense multicell thunderstorm on 14 June 2011. *Wea. Forecasting*, **31**, 827–851, <https://doi.org/10.1175/WAF-D-15-0081.1>.
- Laroche, S., and I. Zawadzki, 1994: A variational analysis method for retrieval of three-dimensional wind field from single-Doppler radar data. *J. Atmos. Sci.*, **51**, 2664–2682, [https://doi.org/10.1175/1520-0469\(1994\)051<2664:AVAMFR>2.0.CO;2](https://doi.org/10.1175/1520-0469(1994)051<2664:AVAMFR>2.0.CO;2).
- Lawless, A. S., S. Gratton, and N. K. Nichols, 2005a: An investigation of incremental 4D-Var using non-tangent linear models. *Quart. J. Roy. Meteor. Soc.*, **131**, 459–476, <https://doi.org/10.1256/qj.04.20>.
- , —, and —, 2005b: Approximate iterative methods for variational data assimilation. *Int. J. Numer. Methods Fluids*, **47**, 1129–1135, <https://doi.org/10.1002/fld.851>.
- Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.*, **22**, 1065–1092, [https://doi.org/10.1175/1520-0450\(1983\)022<1065:BPOTSF>2.0.CO;2](https://doi.org/10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2).
- Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **112**, 1177–1194, <https://doi.org/10.1002/qj.49711247414>.
- Mahale, V. N., G. Zhang, and M. Xue, 2014: Fuzzy logic classification of S-band polarimetric radar echoes to identify three-body scattering and improve data quality. *J. Appl. Meteor. Climatol.*, **53**, 2017–2033, <https://doi.org/10.1175/JAMC-D-13-0358.1>.
- , —, and —, 2016: Characterization of the 14 June 2011 Norman, Oklahoma, downburst through dual-polarization radar observations and hydrometeor classification. *J. Appl. Meteor. Climatol.*, **55**, 2635–2655, <https://doi.org/10.1175/JAMC-D-16-0062.1>.
- Milbrandt, J. A., and M. K. Yau, 2005: A multimoment bulk microphysics parameterization. Part I: Analysis of the role of the spectral shape parameter. *J. Atmos. Sci.*, **62**, 3051–3064, <https://doi.org/10.1175/JAS3534.1>.
- Morrison, H., J. A. Curry, and V. I. Khvorostyanov, 2005: A new double moment microphysics parameterization for application in cloud and climate models. Part I: Description. *J. Atmos. Sci.*, **62**, 1665–1677, <https://doi.org/10.1175/JAS3446.1>.
- NOAA/NCEI, 2011: NEXRAD Level-II radar data. NOAA/NCEI, accessed 20 April 2012, <https://www.ncdc.noaa.gov/data-access/radar-data/nexrad>.
- Park, H. S., A. V. Ryzhkov, D. S. Zrnic, and K.-E. Kim, 2009: The hydrometeor classification algorithm for the polarimetric WSR-88D: Description and application to an MCS. *Wea. Forecasting*, **24**, 730–748, <https://doi.org/10.1175/2008WAF2222205.1>.
- Qiu, C. J., and Q. Xu, 1992: A simple-adjoint method of wind analysis for single-Doppler data. *J. Atmos. Oceanic Technol.*, **9**, 588–598, [https://doi.org/10.1175/1520-0426\(1992\)009<0588:ASAMOW>2.0.CO;2](https://doi.org/10.1175/1520-0426(1992)009<0588:ASAMOW>2.0.CO;2).
- Rodgers, C., 2000: *Inverse Methods for Atmospheric Sounding: Theory and Practice*. Series on Atmospheric, Oceanic, and Planetary Physics, Vol. 2, World Scientific, 256 pp.
- Ryzhkov, A. V., D. S. Zrnic, J. C. Hubbert, V. N. Bringi, J. Vivekanandan, and E. A. Brandes, 2002: Polarimetric radar observations and interpretation of co-cross-polar correlation

- coefficients. *J. Atmos. Oceanic Technol.*, **19**, 340–354, <https://doi.org/10.1175/1520-0426-19.3.340>.
- Sun, J., D. W. Flicker, and D. K. Lilly, 1991: Recovery of three dimensional wind and temperature fields from simulated single Doppler radar data. *J. Atmos. Sci.*, **48**, 876–890, [https://doi.org/10.1175/1520-0469\(1991\)048<0876:ROTDWA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1991)048<0876:ROTDWA>2.0.CO;2).
- Waterman, P. C., 1971: Symmetry, unitarity and geometry in electromagnetic scattering. *Phys. Rev.*, **3**, 825–839, <https://doi.org/10.1103/PHYSREVD.3.825>.
- Yoshikawa, E., V. Chandrasekar, and T. Ushio, 2014: Raindrop size distribution (DSD) retrieval for X-band dual-polarization radar. *J. Atmos. Oceanic Technol.*, **31**, 387–403, <https://doi.org/10.1175/JTECH-D-12-00248.1>.
- Zhang, G., 2015: Comments on “Describing the shape of raindrop size distributions using uncorrelated raindrop mass spectrum parameters.” *J. Appl. Meteor. Climatol.*, **54**, 1970–1976, <https://doi.org/10.1175/JAMC-D-14-0210.1>.
- , 2016: *Weather Radar Polarimetry*. CRC Press, 304 pp.